

Ivan Kassabov

Encyclopaedic Multimedia Dictionary as a Cultural-Linguistic Model

Summary: The present paper reviews the encyclopaedic multimedia dictionary as a result obtained by means of linguistic (lexicographic) modelling. An algorithm was created which tests the model, and the results of the computer processing of the data are more than satisfactory. – The organization of the semantic system of language (with reference to Bulgarian material) is presented in its core and fundamental, necessary and sufficient parts, as well as the cultural-historical system by means of the system of proper names. An encyclopaedic dictionary of this kind may play the role of a cultural-linguistic representator, and from a practical point of view it actually delineates such a model.

Zusammenfassung: Der vorliegende Artikel betrachtet das enzyklopädische Multimedia-Lexikon als Ergebnis linguistischer (lexikographischer) Modellierung. Es wurde ein Algorithmus erarbeitet, um das Modell zu testen, und die Resultate der Datenverarbeitung via Computer sind durchaus befriedigend. – Die Organisation des semantischen Systems der Sprache (bezugnehmend auf bulgarisches Material) wird in seinen zentralen und grundlegenden, notwendigen wie hinreichenden Teilen vorgelegt, wie auch das kulturell-historische System auf Basis der Eigennamen. Ein enzyklopädisches Lexikon dieser Art könnte die Rolle eines kulturell-linguistischen Repräsentators spielen, und aus praktischem Blickwinkel umschreibt es in der Tat ein solches Modell.

When we discuss the possibilities for a complete and (as well simultaneous) economic representation of the human knowledge of the world, we end up with the classical model of the encyclopaedia as a systematic representation of human cultural history (usually from a contemporary point of view). At the same time, every single natural language presents the physical and the cultural universe from the specific viewpoint of the community that speaks this language. Every language segments the continuum of the world in a specific way and organizes its part of content into its own "picture of the world" within the context of the long cultural tradition of the respective civilization.

In contemporary semiotics (following the Stoics, Sextus Empiricus and St. Augustine, and also according to F. de Saussure and L. Hjelmslev, Ju. Lotman and U. Eco) it is considered a genuine truth that language is the primary shaping semiotic system through which all other systems find their material expression. The linguistic "picture of the world" finds its systematic representation in the classical model of the Dictionary. I will adopt the standard view that the Encyclopaedia represents the world objectively and scientifically, while the Dictionary represents the world on popular grounds and naively.

The practical work of compiling a dictionary shows that a pure lexicographic semantization is not possible without an encyclopaedic one. Nor is it possible to achieve a pure encyclopaedic semantization without the lexicographic one. As a matter of fact, it is as rare to find a linguistic dictionary of a "pure" type, as it is to find a "pure" encyclopaedia. They seem to be pure only from a strictly technical point of view, e.g. whether or not they include proper names in their corpora. Strangely enough, however, proper names have to be defined, or else one will need to use his/her encyclopaedic knowledge in order to understand the definition. For example, the word *Bulgarian* is defined as 'a person of the basic population of Bulgaria' and *Sisyphean (toil)* is defined as 'a very hard, but useless work (from the legend of *Sisyphus*)'. On the other hand, both the encyclopaedia and the dictionary contain not only words which are common for the linguistic areas of society and culture, but also words which are names of animals and plants and are indispensable for both the dictionary and the encyclopaedia. However, names of animals or plants are not sufficient by themselves and cannot be described uniquely with the resources of the natural language (as has been

pointed out by A. Wierzbicka in her "Lingua Mentalis" theory). Therefore, it becomes difficult to identify the plants or the animals only by their dictionary definitions. This is why definitions usually add the corresponding Latin name, or the picture of the plant or the animal.

Apart from these clear distinctions between the dictionary and the encyclopaedia, we should be particularly aware of the fact that a "pure" linguistic definition necessarily contains a partial encyclopaedism. Alternatively, "pure" encyclopaedias necessarily contain a partial "linguisticism", if I may say so. This problem has already been pointed out in view of its practical ramifications for the lexicographic theory, but I believe that it is the theory of semantics which has to provide a practical solution. Moreover, in my opinion, the solution lies in the nature and the way the linguistic sign functions (especially in the nature of the word as a fundamental linguistic sign).

From a more practical point of view, this approach motivates the need for the creation of a Universal Dictionary (i.e. an Encyclopaedic *and* a Linguistic one), which results from and at the same time adequately represents the proper combination of the objective-scientific and the linguistic "picture of the world".

My work on an *Encyclopaedic Dictionary of Bulgarian* (in print) applies this particular type of linguistic (lexicographic) modelling. The understanding of language as an organized system calls forth the need for constructing a model for the meaningful part of language, i.e. for its lexico-semantic system. This model can work by itself, i.e. by computer. Such an approach is best revealed through the basic principle which underlies mathematic modelling: an object can be shown to exist truly if and only if there is an algorithm for its description. In other words, in the field of lexicography we can demonstrate the adequacy of the theory by creating a particular model which may serve as the connecting pole between the ideal theory and the real object, i.e. the natural language.

It is well-known that in order for such a model to work, it has to have a finite number of units (no matter how great) and several simple arrangement rules. It is also clear that these units have to be structured in such a way as to create a sufficiently rigid system. The almost unlimited number of units in natural language and their complexity requires that they be artificially limited down to the core part of the lexico-semantic system so

that the model can meet the necessary and sufficient requirements for preserving the specific properties of the object. This limitation derives from the two objective qualities of the description – its completeness and its detailed elaboration. At the same time, the heuristic power of the model, especially when it functions, may compensate for the disadvantages of this limitation. The model shows new regularities and new standard characteristic features of the object, despite the fact that lexical semantics is the most difficult field of modelling. As a first step towards building up a model of the lexico-semantic field, I have constructed a *Minimal Semantics Dictionary* (Sofia 1990) which serves as a theoretical construct representative for the lexical system as a whole. Fig. 1 shows an illustration of the structural organization of the lexico-semantic fields in a network system which reveals the specific “picture of the world” in Bulgarian.

On the basis of this theoretical construct and following the same theoretical principles I have created a *Universal (linguistic-encyclopaedic) Dictionary* which comprises 10 000 units = lexical entries. The latter are organized into 80 terminological fields (with 45 subfields) taking into account the peculiarities of the respective scientific domain. Each of these terminological fields contains only those terms which are necessary for the adequate representation of the respective scientific domain and are at the same time sufficient for the adequate representation of the encyclopaedic knowledge of the world. Each of these terms is defined and they are arranged in alphabetical order. Each lexical entry is divided into three parts – a terminological, a common linguistic, and a metaphorical one. In the first part the entry is defined as a *terminus technicus*, in the second part it is defined with its everyday, naive meaning, and in the third part – with its metaphorical meaning (or the general, non-concrete meaning). As a result of this approach, every word becomes a cross-section of three semantic fields which are principally different from one another: the *terminological-taxonomic*, the *popular-linguistic* and the *associative-metaphorical* field. These fields are domains which belong to three principally different and co-existent “pictures of the world” in a natural language. The *scientific-encyclopaedic* picture presents the scientific knowledge of the current state of the world, perceived as objectively true. The *popular-linguistic* picture presents the common, naive knowledge of the world which has been accumulated over the centuries and is socially significant for the present day.

The *associative-metaphorical* picture is in fact a picture of the possible worlds in the domain of the imaginative, the poetic, and the mythological.

Such a state-of-the-art requires that the semantization and the respective definition of each lexical entry be made to follow each of these three “keys” in accordance with the three respective registers, because the terms, the regular linguistic meanings, and the metaphorical figurative meanings enter different types of semantic relations with the respective meanings of the rest of the words – units of the lexico-semantic system.

The theories of interpretation in the “keys” of the different scientific paradigms, starting with the ones created by Aristoteles and Porphyrius, going through those of the Stoics, Sextus Empiricus, St. Augustine, J. Locke, G. Leibniz, B. Spinoza, the French encyclopaedists (d’Alembert in particular) and including the philosophers of language, were reduced by U. Eco to the well-familiar dichotomy *Dictionary versus Encyclopaedia*. In doing that he represents the principles of semantization by means of the visual-metaphorical juxtaposition between Porphyrius’s “*Tree*”, and a “*Labyrinth*”, arguing at length in favour of the second at the expense of the first.

A closer inspection of the practical problems of interpretation in dictionary definitions, however, suggests that the problem is that both in Dictionaries and Encyclopaedias we interpret not just words or terms, but complex sign units which (even when they appear as whole entry units) act as words (with everyday meanings), as terms (with terminological meanings), and as metaphors (with figurative meanings). The paradox lies in the very urge of lexicographers to introduce “scientificity” and “encyclopaedicity” in their dictionary definitions, and in the attempts of encyclopaedists to offer strict scientific definitions by using the words which are the means of expression in everyday speech. In other words, both the former and the latter try to present themselves not for what they are (by virtue of their own object of study), but for what they are short of being, due to lack of adequate means of interpretation. It is precisely this illusion (which Eco might define as “the illusion of the mirror”) that the suggested Universal (combined) Dictionary-Encyclopaedia helps us overcome.

Thus, we should in fact imagine the semiotic “universe, i.e. the universe of human culture” (as defined by U. Eco) as a “labyrinth of the third (Rhizome) type”. If, however, by the term “*Culture*”, represented by “a semantic encyclopaedia [...which] exists only as a *regulative idea*” rather

than as a finished entity, we mean just to be “able actually to isolate a given portion of the social encyclopaedia so far as it appears useful in order to interpret certain portions of actual discourses (or texts)” (Eco 1984: 84), we exclude the spheres of science and the arts. Actually, this is exactly what d’Alembert suggests in his introduction to the French Encyclopaedia: “The general system of the sciences and the arts is a labyrinth, a meandering road, which the human spirit faces...”

So it turns out that, as regards the “key”, the semantic universe can be represented as a “Tree” (when defining the terminological taxonomies), as a “Labyrinth” of the Meander type (in defining the “purely linguistic” lexical-semantic fields), or as a “Labyrinth-Network” of the Rhizome type (in the semantic representation of the associative fields of the metaphoric and figurative meanings). In reality, however, both the semantic universe and the words – its “atoms” – are, by virtue of their versatile nature, all of the above three. For this reason, in order to achieve an adequate presentation we need such a Universal Dictionary, in which all these aspects are combined.

As a final conclusion, I would like once again to emphasize the historical and the cultural character of the Encyclopaedia and the Dictionary. The Encyclopaedia on its own *differentia specifica* is a presentation of history and culture from a contemporary point of view. The linguistic semantics of the Dictionary contains, transfers and represents the popular (every-day) “*Picture of the world*”. Consequently, the combination of an *Encyclopaedia* and a *Dictionary* is a real *cultural model* achieved by means of a specific type of lexicographic modelling of the semantic domain in order to arrive at a real structural representation of culture as a whole. The best representation of such a semantic universe will be the *Multimedia Encyclopaedia-Dictionary*, which will present and demonstrate the real richness, the internal tension and the flexibility of the relationships and their variety in the organization of such a complex sign-system.

Reference

Eco, Umberto (1984). *Semiotics and the Philosophy of Language*. Bloomington, IN: Indiana University Press

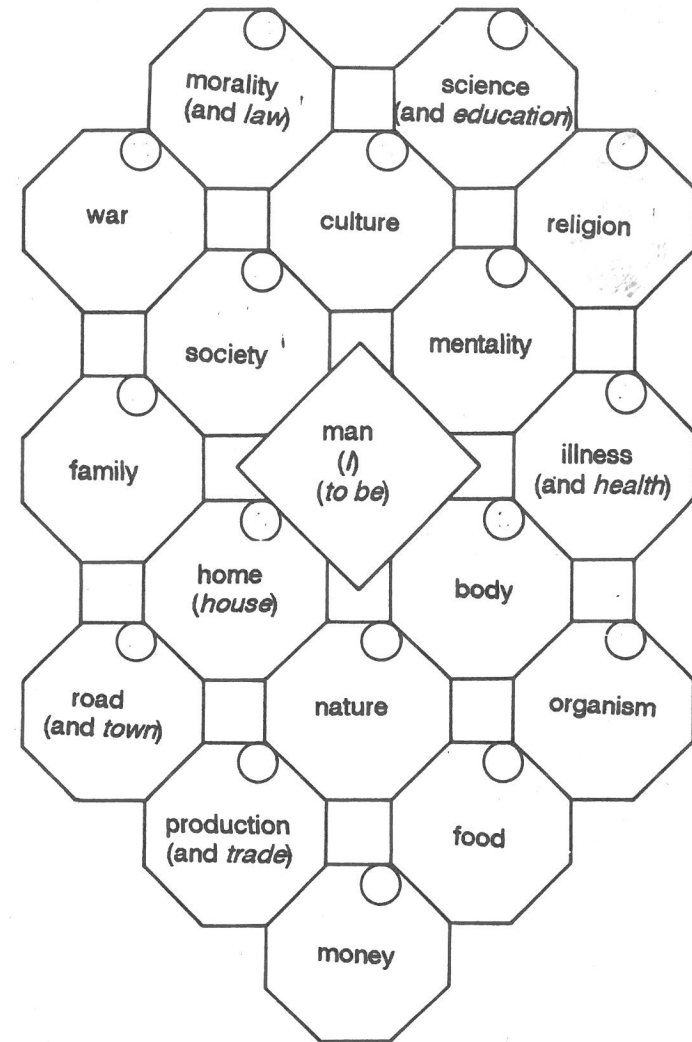


Fig. 1: The Bulgarian “Picture of the World” as a Cultural Model